Supplementary Information

Index of Supplementary Materials

Supplementary Note

NCI DCEG Cancer Genomics Research Laboratory:

Joseph Boland, Laurie Burdett, Belynda Hicks, Herbert Higson, Amy A. Hutchinson, Kristine Jones, Sally Larson, Hyo Jung Lee, Michael Malasky, Michelle Manning, Adri O'Neil, Aurelie Vogt, Rebecca Wasser.

NCI DCEG Cancer Sequencing Working Group

Blanche P. Alter, Neelam Giri, Mark H. Greene, Allan Hildesheim, Nan Hu, Jennifer Loud, Phuong L. Mai, Lisa Mirabello, Lindsay Morton, Dilys Parry, Anand Pathak, Siegal Sadetzki, Douglas R. Stewart, Phil Taylor, Geoffrey Tobias, Jeannette Wong, Guoqin Yu.

French Familial Melanoma Study Group

Caroline Abadie, Bertrand Bachollet, Patrcik Benusiglio, Yves-Jean Bignon, Marie-Noelle Bonnet-Dupeyron, Olivier Caron, Isabelle Coupier, Dominique Courouge- Dorcier, Sophie Dalac, Stéphane Dalle, Sébastien Debarbieux, Capucine Delnatte, Olivier Dereure, Hélène Dreyfus, Catherine Dugast, Sandra Fert-Ferrer, Thierry Frébourg, Marion Gauthier- Villars, Damien Giacchero, Florent Grange, Jean-Jacques Grob, Ewa Hainault, Olivier Ingster, Bruno Labeille, Christine Lasset, Valérie Layet, Marine Lebrun, Dominique Leroux, Jean-Marc Limacher, Michel Longy, Sandrine Mansard, Ludovic Mansuy, Tanguy Martin-Denavit, Christina Mateus, Michèle Mathieu, Eve Maubec, Nicolas Meyer, Gwenaël Nadeau, Laurence Olivier-Faivre, Jean-Luc Perrot, Nicolas Poulhalon, Pascal Pujol, Caroline Robert, Annik Rossi, Luc Thomas, Julie Tinat, Pierre Vabres, Linda Vincent-Fetita, Danièle Zaharia, Hélène Zattara.

Exome Sequencing:

DNA Preparation: For each sample, 500 ng genomic DNA was sheared with a Covaris E210 Sonicator (Covaris, Inc., Woburn, MA, USA) to an average size of 200-300 bp. An adapter-ligated library was prepared with the Bioo Scientific NEXTflex™ Pre-Capture Combo Kit (Bioo Scientific, Austin, TX, USA) according to Bioo-provided protocol.

Pre-Hybridization LM-PCR: Genomic DNA sample libraries were amplified pre-hybridization by ligation-mediated PCR consisting of one reaction containing 20 µL DNA, 16 µL MBG Water, 12 uL Bioo NEXTflex™ PCR Master Mix (Bioo Scientific, Austin, TX, USA), 2 µL NEXTflex Primer Mix, 12.5 µM (includes two primers whose sequences are: 5'-AATGATACGGCGACCACCGAGATCTACAC-3' and 5'-CAAGCAGAAGACGGCATACGAGAT-3'). PCR cycling conditions were as follows: 98˚C for 2 minutes, followed by 5 cycles of 98˚C for 30 s, 60˚C for 30 s, 72˚C for 60 s. The last step was an extension at 72˚C for 4 minutes. The reaction was kept at 4˚C until further processing. The amplified material was cleaned with Agencourt AMPure XP Reagent (Beckman Coulter Inc, Brea, CA, USA) according to the Bioo-provided Protocol. Amplified sample libraries were quantified using Quant-iT™ PicoGreen dsDNA Reagent (Life Technologies, Carlsbad, CA, USA).

Liquid Phase Sequence Capture: Prior to hybridization, 4 amplified sample libraries with unique index adapters were combined in equal amounts (275 ng) into 1.1 µg pools for multiplex sequence capture. Exome sequence capture was performed with NimbleGen's SeqCap EZ Human Exome Library v3.0 (Roche NimbleGen, Inc., Madison, WI, USA). Prior to hybridization the following components were added to the 1.1 ug pooled sample library, 4 µL of NEXTflex HE Universal Oligo 1, 250 µM (5'-AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT-3'), 5 µL of each INV-HE blocking oligo, 50 µM each, complementary to the indexes in the pool (5'-CAAGCAGAAGACGGCATACGAGAT**X**GTGACT GGAGTTCAGACGTGTGCTCTTCCGATCT/C3 Spacer/-3', where **X** is 6-bases of sequence specific to adapter index used for library construction), and 5 µL of 1 mg/mL COT-1 DNA (Invitrogen, Inc., Carlsbad, CA, USA). Samples were dried down by puncturing a hole in plate seal and processing in an Eppendorf 5301 Vacuum Concentrator (Eppendorf, Hauppauge, NY, USA) set to 60˚C for approximately 1 hour. To each dried pool, 7.5 µL of NimbleGen Hybridization Buffer and 3.0 µL of NimbleGen Hybridization Component A were added, and placed in a heating block for 10 minutes at 95˚C. The mixture was then transferred to 4.5 µL of EZ Exome Probe Library and hybridized at 47˚C for 64 to 72 hours. Washing and recovery of captured DNA were performed as described in NimbleGen SeqCap EZ Library SR Protocol.

Post-Hybridization LM-PCR: Pools of captured DNA were amplified by ligation-mediated PCR consisting of one reaction for each pool containing 20µl captured DNA, 26µl MGB Water, 50µl NEXTflex™ LM-PCR Master Mix (Bioo Scientific, Austin, TX, USA), 2µl NEXTflex™ PCR Oligo 1 (100µM, 5'-AATGATACGGCGACCACCGAGATCTACAC-3') and 2µl NEXTflex™ PCR Oligo 2 (100µM, 5'-CAAGCAGAAGACGGCATACGAGAT-3'). Post-hybridization amplification cycling conditions were as follows: 98˚C for 30 s, followed by 10 cycles of 98˚C for 10 s, 65˚C for 30 s, 72˚C for 30 s. The last step was an extension at 72˚C for 5 minutes. The reaction was kept at 4˚C until further

processing. The amplified material was cleaned with Agencourt AMPure XP Reagent (Beckman Coulter Inc, Brea, CA, USA) according to NimbleGen SeqCap EZ Library SR Protocol. Pools of amplified captured DNA were then quantified via Kapa's Library Quantification Kit for Illumina (Kapa Biosystems, Woburn, MA, USA) on the LightCycler 480 (Roche, Indianapolis, IN, USA).

Sequencing: The resulting post-capture enriched multiplexed sequencing libraries were diluted to 15 pM and used in cluster formation on an Illumina cBOT (Illumina, San Diego, CA, USA) and paired-end sequencing was performed using an Illumina HiSeq following Illumina-provided protocols for 2x100 paired-end sequencing. Each exome was sequenced to high-depth in order to achieve a minimum threshold of 80% of coding sequence (CDS) bases covered with at least 15 reads, based on the UCSC hg19 "known gene" transcripts (http://genome.ucsc.edu/). An average CDS coverage of over 30-fold was typically required to meet this minimum threshold.

Bioinformatic Analysis:

The human reference genome and the "known gene" transcript annotation were downloaded from the UCSC database (http://genome.ucsc.edu/), version hg19 (corresponding to Genome Reference Consortium assembly GRCh37). Sequencing reads were first trimmed using the Trimmomatic program[1], which marked all low-quality stretches (average < Q15 in a 4-bp sliding window) and reported the longest high-quality stretch of each read. Only read pairs with both ends no shorter than 36 bp were used. Reads were aligned to the hg19 reference genome using the Novoalign software version 2.07.14 (http://www.novocraft.com). Duplicate reads based on paired ends aligning to the same start locations due to either optical or PCR artifacts were removed from further analysis using the MarkDuplicates module of the Picard software (http://picard.sourceforge.net/). Additionally, our analysis used only properly aligned read pairs, in the sense that the two ends of each pair must be mapped to the reference in complementary directions and must reflect a reasonable fragment length (300+/-100 bp). These high-quality alignments for each individual were refined using a local realignment strategy around known and novel sites of insertion and deletion polymorphisms using the RealignerTargetCreator and IndelRealigner modules from the Genome Analysis Toolkit (GATK)[2]. Variant discovery and genotype calling of multi-allelic substitutions, insertions and deletions was performed on all individuals globally using the UnifiedGenotyper module from GATK with the minimum call quality parameter set to 30, with necessary recalibrations of quality scores and variant quality scores following the "best practice" recommended by the GATK development team. Annotation, fitting genetic models, and filtering of each variant locus was performed using a locally-developed custom software pipeline which took use of public data from the UCSC GoldenPath database (http://hgdownload.cse.ucsc.edu/goldenPath/hg19/database/), the ESP6500 dataset from University of Washington's Exome Sequencing Project (http://evs.gs.washington.edu/EVS/), the Institute of Systems Biology KAVIAR (Known VARiants) database[3], the National Center for Biotechnology Information dbSNP database[4] build 137, and the 1000 Genomes Project[5].

Chromosome-based exact test

Currently available methods for association testing in related subjects rely on asymptotic approximations to calculate p-values, which is not suitable for rare variants. Therefore, we developed a chromosome-based test to measure the association between CMM and the Ser270Asn variant by appropriately accounting for the relatedness for familial CMM cases within the same families. We first conducted kinship analysis, which did not reveal cryptic relatedness across families using a kinship coefficient cutoff 0.05. We then inferred IBD at the Ser270Asn locus using SNP genotypes based on exome sequencing data to identify independent alleles and the number of patients carrying each allele in each family. We performed whole exome sequencing in 101 familial subjects, including 97 CMM cases and four obligate carriers (assuming dominant inheritance). In the 97 familial CMM cases, we identified 153 independent chromosomes at the locus, among which two chromosomes were carried by three related patients, 32 chromosomes were carried by two related patients, and 119 chromosomes were carried by only one patient. Since all genetically unrelated subjects have two independent chromosomes, there were 4,229 independent alleles in total at the locus (153 in families +2,038×2 =4,076 in controls).

In our observed data, 11 patients (denoted as $\boldsymbol{M_0}$) from five unrelated families carried the Ser270Asn variant allele (we excluded the obligate gene carrier from this analysis). Intuitively, a larger value of $\boldsymbol{M_0}$ provides stronger evidence to reject the null hypothesis $\boldsymbol{H_0}$ of no association. In theory, the significance of $\boldsymbol{M_0}$ can be evaluated by permutations in which we would randomly permute the 4,229 chromosomes, count the number of patients (denoted as $\boldsymbol{M}$) carrying the Ser270Asn variant allele in each permutation, and calculate the proportion of permutations with $\boldsymbol{M}\geq 11$. Because chromosomes are independent within families and permutations are performed conditioning on the IBD sharing structure, this procedure automatically accounts for the relatedness among patients in each family. Although the permutation procedure is conceptually straightforward, it is computationally prohibitive for very small p-values. Thus, we performed an exact test by directly calculating the probability $P(\boldsymbol{M}\geq \boldsymbol{M_0}|\boldsymbol{H_0})$.

In this test, we randomly chose five chromosomes (from N=4,229 chromosomes), assigned them as Ser270Asn variant alleles, and counted the number of patients carrying the variant alleles (denoted as $\boldsymbol{M}$). The p-value of the chromosomes-based exact test is defined as the probability $P(\boldsymbol{M}\geq \boldsymbol{M_0}|\boldsymbol{H_0})$. If $\boldsymbol{E_i}=\{i$ Ser270Asn variant alleles assigned to the patient group$\}$ for $i=0,1,\cdots,5$, then

$$P(\boldsymbol{M}\geq \boldsymbol{M_0}=11|\boldsymbol{H_0})=\sum_{i\leq 5} P(\boldsymbol{M}\geq 11,\boldsymbol{E_i}|\boldsymbol{H_0})= \sum_{i\leq 5} P(\boldsymbol{M}\geq 11|\boldsymbol{H_0},\boldsymbol{E_i})\ \boldsymbol{P}(\boldsymbol{E_i}|\boldsymbol{H_0}).$$

Given the IBD sharing pattern, if less than five Ser270Asn alleles are assigned to the case group, the maximum number of patients carrying the allele would be $3+3+2+2=10<11$. Thus, $P(\boldsymbol{M}\geq 11|\boldsymbol{H_0},\boldsymbol{E_i})=0$ for $i=0,1,\cdots,4$. Therefore,

$$P(\boldsymbol{M}\geq \boldsymbol{M_0}=11|\boldsymbol{H_0})= P(\boldsymbol{M}\geq 11|\boldsymbol{H_0},\boldsymbol{E_5})\ \boldsymbol{P}(\boldsymbol{E_5}|\boldsymbol{H_0}).$$

The probability $P(\boldsymbol{E_5}|\boldsymbol{H_0})$ can be calculated using the hypergeometric distribution

$$P(\boldsymbol{E_5}|\boldsymbol{H_0})= \frac{\binom{153}{5}\binom{4,076}{0}}{\binom{4,229}{5}}=5.8\times10^{-8}.$$

Let $(x_1,\cdots,x_5)$ represent the number of patients carrying the five Ser270Asn alleles (unordered). Given the fact that all five Ser270Asn alleles are assigned to the case group, only three combinations of $(x_1,\cdots,x_5)$ allow $\boldsymbol{M}\geq 11$: (3,3,2,2,2), (3,3,2,2,1) and (3,2,2,2,2). Thus,

$$P(\boldsymbol{M}\geq 11|\boldsymbol{H_0},\boldsymbol{E_5})= \frac{\binom{32}{3}+\binom{32}{2}\binom{119}{1}+\binom{2}{1}\binom{32}{4}}{\binom{153}{5}}=2.0\times10^{-4}.$$

Combining the two probabilities gives the final p-value $P(\boldsymbol{M}\geq 11|\boldsymbol{H_0})= 5.8\times10^{-8}\times2.0\times10^{-4}=1.2\times10^{-11}$.

TaqMan® SNP Genotyping:

TaqMan® assay (Life Technologies, Grand Island, NY) validation and SNP genotyping was performed at the Cancer Genomics Research Laboratory, NCI (http://cgf.nci.nih.gov/). TaqMan® employs the 5' nuclease allelic discrimination assay which includes a forward target-specific PCR primer, reverse primer, and TaqMan® MGB probes labeled with two dyes (FAM and VIC) or (FAM and TAMRA).

Assay Validation: Loci that produced successful TaqMan® assay designs were validated using the 102 Variant GPS (http://variantgps.nci.nih.gov/) population panel and the 262 HapMap control panel (Coriell Cell Repositories, Camden, NJ). Assay validation was based on concordance with reported HapMap genotypes (when available). Frequency information is reported using the 102 and 262 validation sets, available from the Variant GPS website.

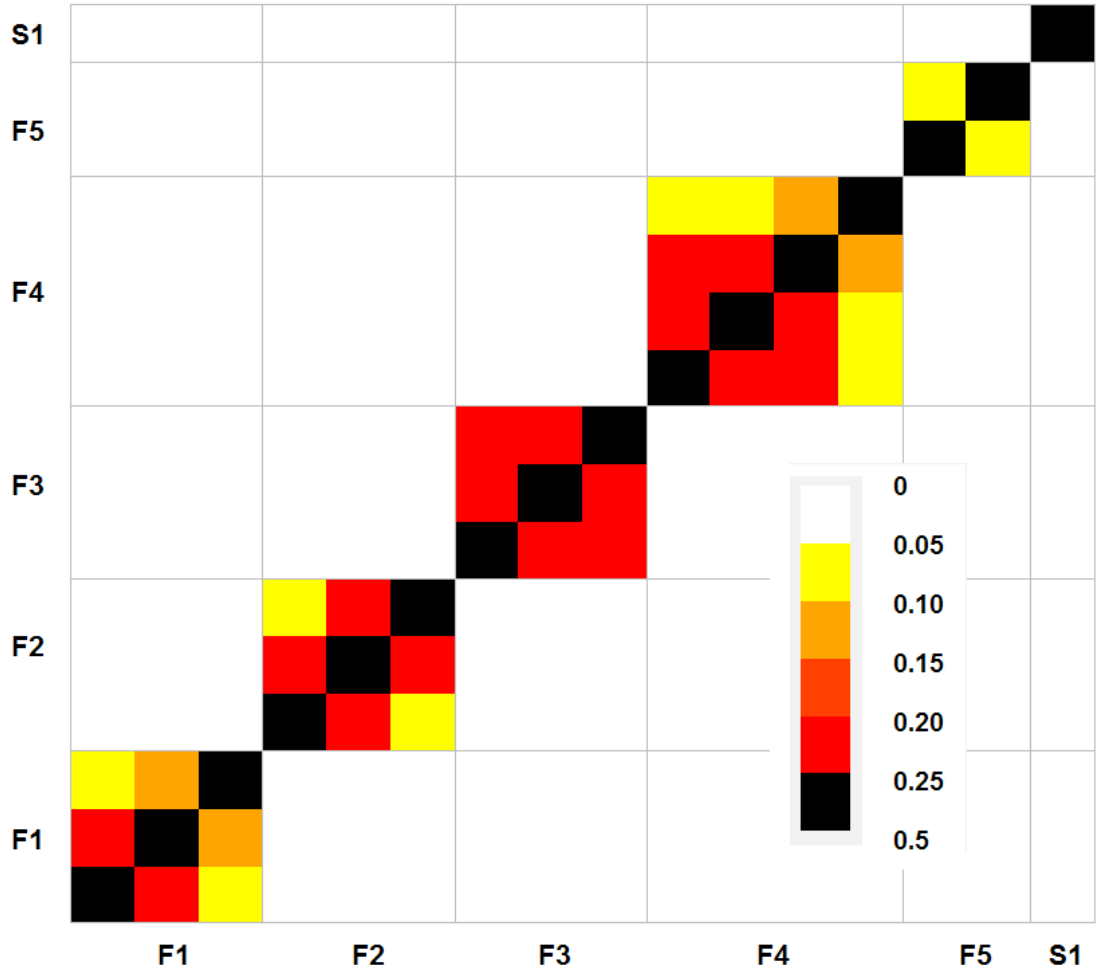
SNP Genotyping: 5 ng of sample DNA, according to Quant-iT™ PicoGreen dsDNA quantitation (Life Technologies), was transferred into 384-well plates (Life Technologies) and dried down. Additionally, 5 ng of assay-specific controls, based on the validation results, were applied to pre-determined wells of the assay plates to guide analysis and overall quality and 5 ng of universal internal controls (NA07057 and NTC) were added to random locations of 384-well plates to provide a unique fingerprint for each plate for overall quality assurance.

Genotyping was performed using 5 uL reaction volumes consisting of: 2.5 uL of 2X KAPA Probe Fast MasterMix (Kapa Biosystems, Cape Town, South Africa), 0.25 uL of 20X TaqMan® assay-specific mix of primers and probes, and 2.25 uL of MBG Water. Plates were heat-sealed using diamond optical seals (Thermo Scientific, Waltham, MA) and PCR was performed using 9700 Thermal Cycler (Life Technologies) with the following conditions: 95˚C hold for 3 min, 40 cycles of 95˚C for 3 sec and X˚C for 30 sec (where X is the optimized annealing temperature listed), and 10˚C hold.

Endpoint reads were evaluated using the 7900HT Sequence Detection System (Life Technologies) and cluster analysis was performed using SDS v2.2.2 software (Life Technologies). Cluster analysis is performed using the Allelic Discrimination Plot, which is an X-Y scatter plot of FAM and VIC dyes, containing four distinct clusters which represent three possible genotypes: Allele 1 Homozygous (Y-Axis), Allele 2 Homozygous (X-Axis), and Allele 1/Allele 2 Heterozygous (Diagonal Axis) while the fourth cluster is at the origin and contains no amplification (NTCs). Analyzed data was imported into a LIMS where concordance of assay-specific genotyping controls and internal controls were confirmed.

References:

1. Lohse, M. *et al.* RobiNA: a user-friendly, integrated software solution for RNA-Seq-based transcriptomics. *Nucleic Acids Res* **40**, W622-7 (2012).
2. DePristo, M.A. *et al.* A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* **43**, 491-8 (2011).
3. Glusman, G., Caballero, J., Mauldin, D.E., Hood, L. & Roach, J.C. Kaviar: an accessible system for testing SNV novelty. *Bioinformatics* **27**, 3216-7 (2011).
4. Sherry, S.T. *et al.* dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res* **29**, 308-11 (2001).
5. Abecasis, G.R. *et al.* A map of human genome variation from population-scale sequencing. *Nature* **467**, 1061-73 (2010).

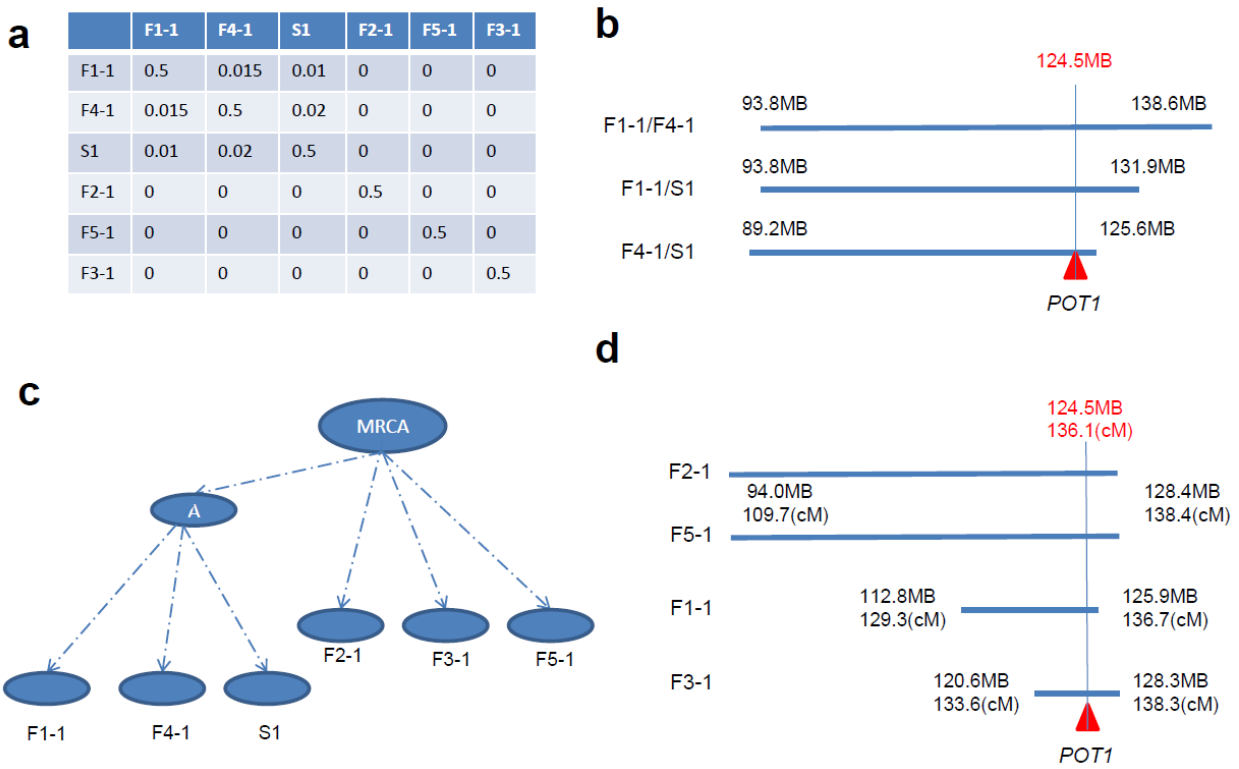


Supplementary Figure 1. Heat map of the estimated kinship coefficients for *POT1*-Ser270Asn variant carriers

POT1-Ser270Asn variant carriers (from Italian families (F1-F5) plus a sporadic case (S1)) and Environment And Genetics in Lung cancer Etiology (EAGLE) controls were genotyped using the Illumina HumanOmniExpress BeadChip (>700,000 common SNPs) and the HumanHap550 BeadChip (~550,000 SNPs), respectively. 306,684 autosomal SNPs that were on both platforms were used for the estimation of kinship coefficients using PLINK (http://pngu.mgh.harvard.edu/~purcell/plink/, "--genome" option). Each block represents a family. None of the subject pairs from different families had an estimated kinship coefficient greater than 0.05.

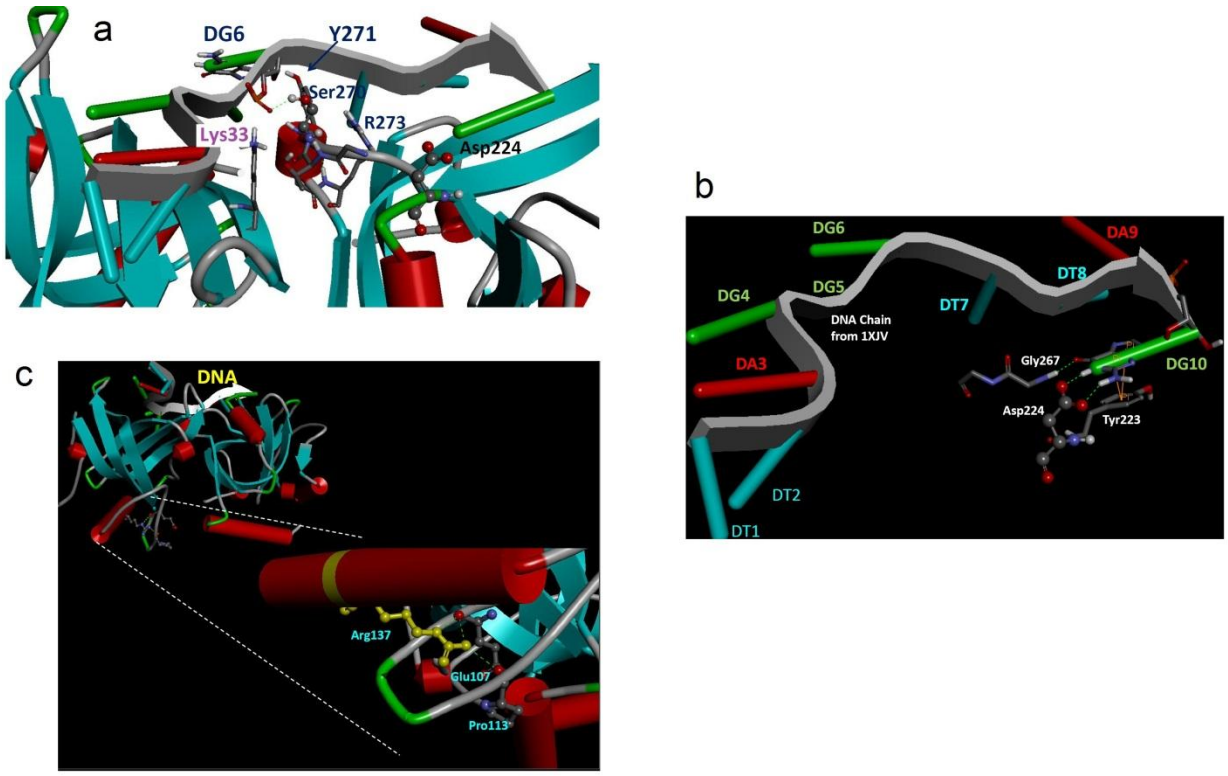
a

	F1-1	F4-1	S1	F2-1	F5-1	F3-1
F1-1	0.5	0.015	0.01	0	0	0
F4-1	0.015	0.5	0.02	0	0	0
S1	0.01	0.02	0.5	0	0	0
F2-1	0	0	0	0.5	0	0
F5-1	0	0	0	0	0.5	0
F3-1	0	0	0	0	0	0.5



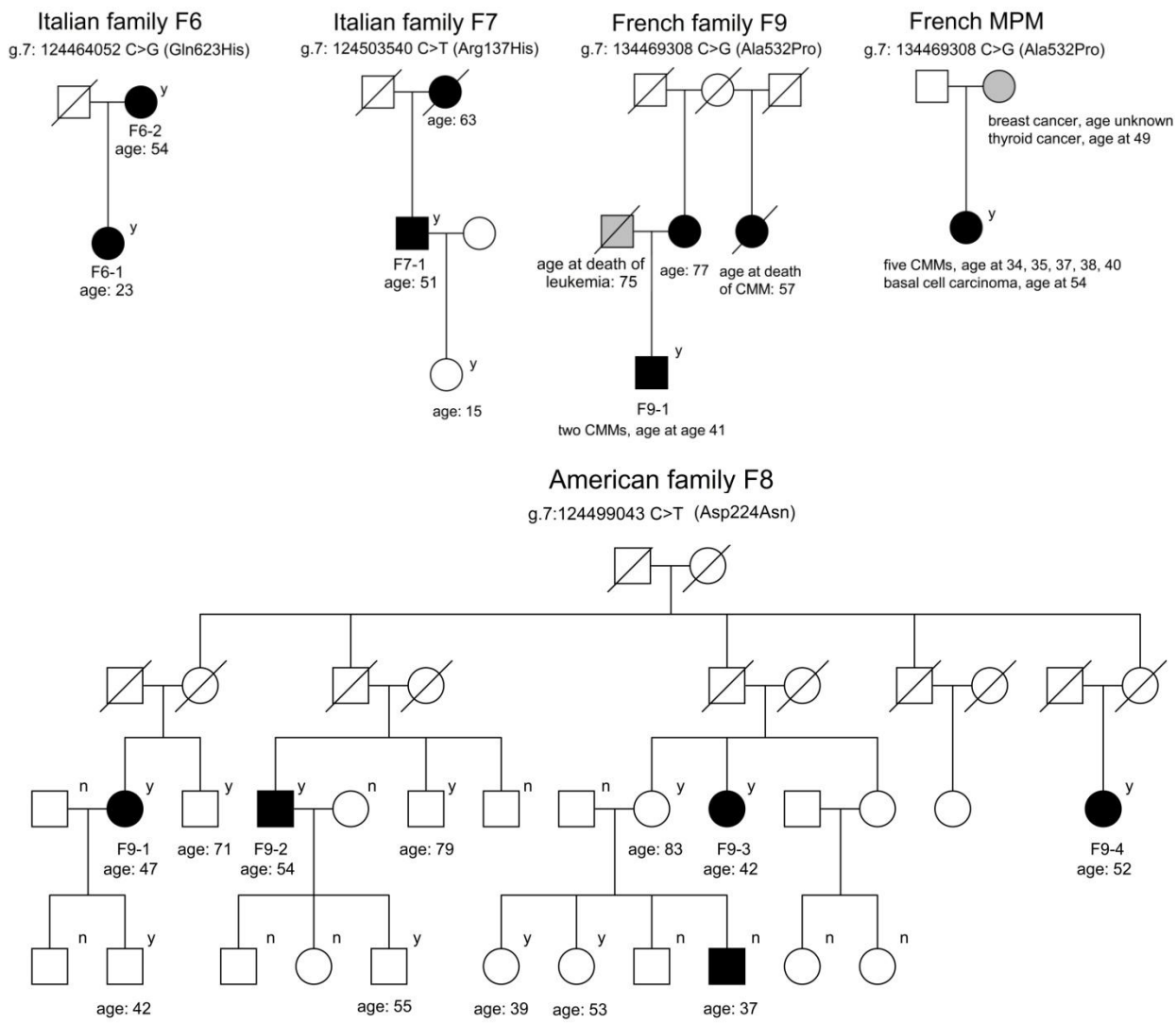
Supplementary Figure 2. Estimation of the age of the most recent common ancestor (MRCA) carrying the Ser270Asn variant.

All analyses were based on the five probands from the Italian CMM families (F1-1,F2-1,F3-1,F4-1,F5-1) and the one sporadic case (S1) who carried the Ser270Asn variant. (**a**) Estimated kinship coefficients for Ser270Asn variant carriers. F1-1, F4-1 and S1 show weak correlation. (**b**) Subjects F1-1, F4-1 and S1 shared long haplotypes encompassing the *POT1* region (horizontal blue lines). GERMLINE analysis (http://www.cs.columbia.edu/~gusev/germline/) revealed that multiple additional genomic segments were also shared among these three subjects (data not shown) suggesting that they descended from a common ancestor. The megabase location is provided for the shared haplotype boundaries and for *POT1*. The red triangle shows the *POT1* location at position 124.5MB based on hg19. (**c**) The lineage used to estimate the age of the MRCA carrying the Ser270Asn variant. From panels a and b, F1-1, F4-1 and S1 descended from a common ancestor, designated as A. A, F2-1, F3-1 and F5-1 descended independently from the MRCA. (**d**) To estimate the age of the MRCA, we used the independent haplotypes from subjects F2-1, F3-1, F5-1 and from F1-1, which we used to represent the subjects who descended from common ancestor A. The genetic distance (numbers with cM, centimorgan, in parentheses) and megabase location for the boundaries of the haplotypes and *POT1* (in red) are also provided.



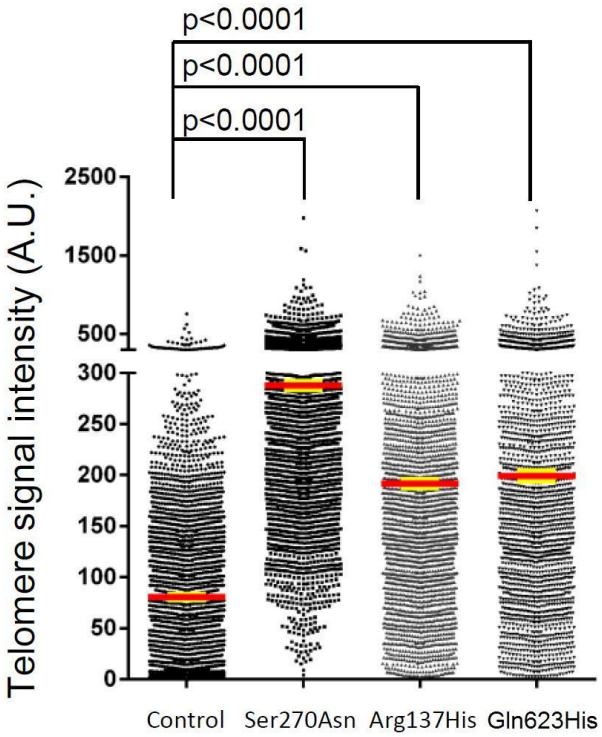
Supplementary Figure 3. Structural illustrations of *POT1* rare variants in the N-terminal OB domains.

(a) Ser270 is an exposed residue and from the crystal structure (1XJV) it is evident that it interacts with the phosphate (PI) backbone of nucleotide guanine 6 (DG6) of the telomeric single-stranded DNA (ssDNA) using H-bonds (dotted green lines). Important residues surrounding Ser270 (ball-and-stick) are shown in stick form. **(b)** Asp224 may form coordinated interactions with Gly267, which mimic the role of a guanine-cytosine base-pair interactions and appear to be further strengthened by the favorable positioning of Tyr223, which forms energetically favorable PI-PI interactions (orange lines) with DG10 from the ssDNA (PDB: 1XJV). **(c)** Key residues surrounding Arg137 (highlighted in yellow) are shown in ball-and-stick mode. Arg137 is involved in H-bond interactions (dotted green lines) with Pro113 and Glu107 and also the possibility of forming a salt-bridge with the latter. These concerted interactions could play a key role in the structural integrity of the fold and indirectly influence the DNA binding when mutated. Research Collaboratory for Structural Bioinformatics (RCSB) Protein Data Bank (PDB) ID 1XJV was used for highlighting the interactions.



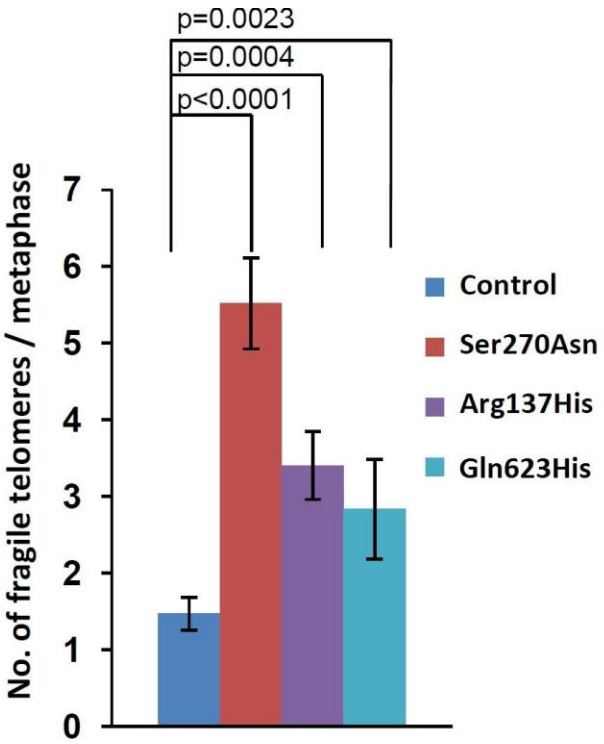
Supplementary Fig 4. Pedigrees of cutaneous malignant melanoma (CMM)-prone families with rare variants in *POT1*.

Solid squares and circles: CMM cases; Grey: other cancers; Circles: females; Squares: males. "y" indicates a mutation carrier and "n" indicates a non-carrier. "age" is the age at diagnosis for CMM cases and age at exam for unaffected family members.



Supplementary Figure 5. Telomere length in *ex vivo* stimulated PBMCs from individuals with *POT1* variants.

Telomere signal intensity reflecting telomere length was quantified by quantitative telomere-fluorescence in situ hybridization (Q-FISH) in *ex vivo* stimulated PMBCs from *POT1*-variant carriers (N=4, including two with Ser270Asn, and one each with Arg137His and Gln623His) and controls (age-matched melanoma cases without *POT1* variants, N=3). Bars (in red) denote mean telomere signal intensity depicted in arbitrary units (A.U) in each group. Error bars (in yellow) indicate standard deviation of errors. *P* values were obtained from two-sided Student's *t*-test.



Supplementary Figure 6. Incidence of fragile telomeres in *ex vivo* stimulated PBMCs from individuals with *POT1* variants.

Numbers of fragile telomeres per metaphase spread of *ex vivo* stimulated PMBCs from *POT1*-variant carriers (N=4, including two with Ser270Asn, and one each with Arg137His and Gln623His) and controls (age-matched melanoma cases without *POT1* variants, N=3). At least 30 metaphases/sample were counted. Error bars indicate standard deviation. *P* values were obtained from two-sided Student's *t*-test.

Supplementary Table 1. Clinical characteristics of melanoma patients with *POT1* mutations

Mutation	Country	Familial Sporadic MPM	Subject_ID	Case Status	Age[a]	Sex	N. melanomas	Presence of DN
Ser270Asn	ITALY	Familial	F1-1	Affected	15	F	2	YES
Ser270Asn	ITALY	Familial	F1-2	Affected	56	M	1	YES
Ser270Asn	ITALY	Familial	F1-3	Obligate Carrier	64	M	0	YES
Ser270Asn	ITALY	Familial	F2-1	Affected	18	M	1	IND[b]
Ser270Asn	ITALY	Familial	F2-2	Affected	55	F	2	YES
Ser270Asn	ITALY	Familial	F3-1	Affected	22	M	1	YES
Ser270Asn	ITALY	Familial	F3-2	Affected	22	F	1	YES
Ser270Asn	ITALY	Familial	F3-3	Affected	53	M	1	YES
Ser270Asn	ITALY	Familial	F4-1	Affected	39	F	1	YES
Ser270Asn	ITALY	Familial	F4-2	Affected	37	M	1	NO
Ser270Asn	ITALY	Familial	F4-3	Affected	66	M	1	YES
Ser270Asn	ITALY	Familial	F5-1	Affected	51	M	1	NO
Ser270Asn	ITALY	Sporadic	S-1	Affected	43	F	1	YES
Gln623His	ITALY	Familial	F6-1	Affected	54	F	1	YES
Gln623His	ITALY	Familial	F6-2	Affected	23	F	1	YES
Arg137His	ITALY	Familial	F7-1	Affected	51	M	1	YES
Asp224Asn	USA	Familial	F8-1	Affected	47	F	1	YES
Asp224Asn	USA	Familial	F8-2	Affected	54	M	1	IND[b]
Asp224Asn	USA	Familial	F8-3	Affected	42	F	1	YES
Asp224Asn	USA	Familial	F8-4	Affected	52	F	1	YES
Asp224Asn	ITALY	Sporadic	S-5	Affected	55	M	2	NO
Ala532Pro	FRANCE	Familial	F9-1	Affected	41	M	2	NO
Ala532Pro	FRANCE	Sporadic	MPM	Affected	34	F	5	NO

[a]Age at diagnosis of melanoma
[b]Indeterminate

Supplementary Table 2: *In silico* analysis of *POT1* rare variants identified in families.

Variant[a]	**Arg137His**	**Asp224Asn**	**Ser270Asn**	**Ala532Pro**	**Gln623His**
Genomic location (Chromosome 7)[b]	g.124503540 C>T	g.124499043 C>T	g.124493086 C>T	g.124469308 C>G	g.124464052 C>G
Codon change[c]	CGT>CaT	GAT>aAT	AGT>AaT	GCA>cCA	CAG>CAc
In silico predictions					
SIFT	Damaging	Damaging	Damaging	Damaging	Damaging
Provean	Deleterious	Deleterious	Neutral	Neutral	Neutral
PolyPhen-2[d]	Probably damaging	Probably damaging	Benign	Possibly damaging	Probably damaging
MutationTaster	Disease causing; May affect protein feature; Potential splice site changes	Disease causing	Disease causing; May affect protein feature; Potential splice site changes	Polymorphism; Potential splice site change	Disease causing; Potential splice site changes
MutationAssessor	Medium functional impact	Medium functional impact	Medium functional impact	Low functional impact	Medium functional impact

[a]These amino acid changes are based on NM_015450/NP_056265.
[b]Annotation based on the reference human genome UCSC build hg19/Genome Reference Consortium GRCh37.
[c]Lower case letter indicates the mutant nucleotide.
[d]Predictions were made based on the HumDiv model.

Supplementary Table 3. Rare variants[a] identified from targeted sequencing of *POT1* in 768 melanoma cases and 768 controls in Italy.

Chr position	Ref	Var	#Cases	#Controls	Gene location
124463996	T	C	1	0	UTR,UTR3
124464153	C	A	0	1	intron
124464154	G	A	0	1	intron
124465228	A	G	0	1	intron
124465427	T	G	1	0	intron
124465480	C	A	1	0	intron
124465487	C	A	0	1	intron
124465501	T	C	1	0	intron
124467290	A	G	1	4	CDS,UTR
124469270	A	G	0	1	intron
124481033	T	C	1	0	CDS,UTR
124482897	T	C	5	2	CDS,UTR
124486966	G	C	1	1	intron
124491835	G	C	0	1	intron
124491845	G	A	1	0	intron
124491857	A	T	2	1	intron
124491883	A	C	0	1	intron
124491996	T	C	2	0	CDS,UTR
124492979	G	A	0	1	intron
124493026	T	A	1	0	CDS,UTR
124493058	T	G	1	1	CDS,UTR
124493086	C	T	2	0	CDS,UTR
124498981	A	G	2	0	intron
124499036	T	G	2	2	CDS,UTR
124499044	G	A	2	0	CDS,UTR
124499212	G	T	4	1	intron

124503469	A	T	0	1	CDS,UTR
124503608	G	C	2	0	CDS,UTR,UTR5
124532305	A	C	2	0	intron
124532425	T	C	0	1	CDS,UTR,UTR5
124532434	C	A	1	0	CDS,UTR,UTR5
124532487	A	T	0	1	intron
124537288	C	T	2	0	intron

[a]Rare variants are defined as <1% in ESP and 1000 Genome.

Supplementary Table 4: Other potentially important rare *POT1* variants identified in a single French familial or MPM (multiple primary melanoma) case.

Variant[a]	Ile78Thr	Phe106Ser	Ile114Asn	Ser421Leu	Leu454Ser	Ile455Val	Intronic	Intronic	Intronic
Genomic location (Chromosome 7)[b]	g.124510987 A>G	g.124503633 A>G	g.124503609 A>T	g.124481134 G>A	g.124481035 A>G	g.124481033 T>C	g.124537326 C>G	g.124467236 T>C	g.124467225 T>C
Codon change[c]	ATT>AcT	TTT>TcT	ATC>AaC	TCA>TtA	TTG>TcG	ATA>gTA	N/A	N/A	N/A
Melanoma case	MPM	MPM	Familial	Familial	MPM	Familial	MPM	MPM	MPM
In silico predictions									
SIFT	Damaging	Damaging	Damaging	Damaging	Damaging	Damaging	N/A	N/A	N/A
Provean	Deleterious	Deleterious	Deleterious	Deleterious	Neutral	Neutral	N/A	N/A	N/A
PolyPhen-2[d]	Probably damaging	Probably damaging	Possibly damaging	Benign	Probably damaging	Possibly damaging	N/A	N/A	N/A
MutationTaster	Disease causing; May affect protein features; Potential splice site changes	Disease causing; May affect protein features; Potential splice site changes	Disease causing; May affect protein features; Potential splice site changes	Disease causing	Disease causing	Disease causing; May affect protein features; Potential splice site changes	May affect protein features; Potential splice site change	Potential splice site change	Potential splice site change
MutationAssessor	Medium functional impact	Medium functional impact	Low functional impact	Medium functional impact	Medium functional impact	Medium functional impact	N/A	N/A	N/A
GERP[e]	4.71	5.44	5.44	5.55	5.41	5.41	0.229	3.92	-7.2
PhyloP[e]	2.023	3.333	3.333	4.031	3.301	3.333	0.043	1.497	-0.877
PhastCons[e]	1	1	0.943	0.995	1	1	0	0.09	0

[a]These amino acid changes are based on NM_015450/NP_056265.

[b]Annotation based on the reference human genome UCSC build hg19/Genome Reference Consortium GRCh37.

[c]Lower case letter indicates the mutant nucleotide.

[d]Predictions were made based on the HumDiv model.

[e]GERP, PhyloP, and phastCons are programs to calculate conservation scores. GERP identifies constrained elements in multiple alignments by quantifying substitution deficits. GERP value > 3 indicates that a site is under highly evolutionary constraint. PhastCons and phyloP are both methods to determine the grade of conservation of a given nucleotide. PhastCons values vary between 0 and 1, the closer the value is to 1, the more probable the nucleotide is conserved. PhyloP values vary between -14 and +6. Sites predicted to be conserved are assigned positive scores, while sites predicted to be fast-evolving are assigned negative scores.